Indonesian Journal of Computer Science
ISSN 2549-7286 (*online*)
Jln. Khatib Sulaiman Dalam No. 1, Padang, Indonesia
Website: ijcs.stmikindonesia.ac.id | E-mail: ijcs@stmikindonesia.ac.id

Bitcoin Price Prediction Using Hybrid LSTM-GRU Models

Nashwan M. Salih[1], Adnan M. Abdulazeez[2]
Nashwan.hussein@dpu.edu.krd, adnan.mohsin@dpu.edu.krd
[1]CIS Dept., Zakho Technical College, Duhok Polytechnic University, Duhok, Iraq
[1]IT Dept., Technical College of Informatics, Akre University for Applied Sciences, Duhok, Iraq
[2]IT Dept., Duhok Technical College, Duhok Polytechnic University, Duhok, Iraq

Article Information

Submitted : 23 Jan 2024
Reviewed: 26 Jan 2024
Accepted : 10 Feb 2024

Keywords

Bitcoin Price Prediction, Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Deep Learning.

Abstract

Cryptocurrency price prediction is a challenging task due to the inherent volatility and complexity of the market. In this research, we propose a hybrid Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) neural network model for predicting Bitcoin prices. The model is implemented using the TensorFlow and Keras libraries and is evaluated on historical Bitcoin price data obtained from Yahoo Finance. Our approach aims to leverage the strengths of both LSTM and GRU architectures to enhance the accuracy of price predictions. The results suggest that the proposed hybrid LSTM-GRU model holds promise for effectively capturing the complex dynamics of cryptocurrency markets, addressing the challenges associated with traditional time-series analysis techniques.

A. Introduction

Cryptocurrencies, led by Bitcoin, have surged into global prominence, reshaping financial landscapes and garnering attention from investors, traders, and researchers alike. The decentralized nature and volatility of these digital assets have created a burgeoning interest in predicting their prices, aiming to unlock insights into their intricate market dynamics. Among the numerous methods employed for price forecasting, artificial intelligence and deep learning models have emerged as promising tools, offering predictive capabilities that leverage historical data and intricate patterns [1].

Bitcoin, as the pioneer of cryptocurrencies, remains at the epicenter of this paradigm shift in finance. Since its inception by the pseudonymous Satoshi Nakamoto in 2009, Bitcoin has evolved from an obscure digital concept to a sought-after asset, captivating both institutional investors and retail traders. Its price volatility, characterized by drastic fluctuations and surges, has presented both opportunities and challenges for market participants, stimulating intense research to discern and forecast its price movements [2].

This study delves into the realm of predictive analytics by exploring the efficacy of a hybrid Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) neural network model in forecasting Bitcoin prices. LSTM and GRU, both variants of recurrent neural networks (RNNs), possess unique architectures designed to capture temporal dependencies within sequential data, making them well-suited for time series prediction tasks. The hybridization of these models aims to combine their strengths, potentially enhancing the predictive performance and robustness when applied to Bitcoin price data [3][4].

The primary objective of this research is to investigate the feasibility and effectiveness of a hybrid LSTM-GRU model in predicting Bitcoin prices using historical data sourced from Yahoo Finance. The prediction of Bitcoin prices holds multifaceted significance in various domains, including finance, investment strategy formulation, risk management, and cryptocurrency market analysis. Accurate price forecasting facilitates informed decision-making for traders, investors, and financial institutions, enabling them to optimize their investment portfolios, hedge risks, and capitalize on market opportunities [5].

Furthermore, a successful implementation of a hybrid LSTM- GRU model for Bitcoin price prediction not only contributes to advancements in predictive modeling but also sheds light on the underlying behavioral dynamics of cryptocurrency markets. Understanding and predicting the price movements of Bitcoin can potentially unlock insights into the broader crypto market sentiment, adoption trends, and macroeconomic factors influencing its valuation.

B. Related works

Seabe et al. [4] suggested that prices of cryptocurrencies were predicted using three different deep learning techniques: LSTM, GRU, and Bi-LSTM. Two metrics, RMSE and MAPE, were used to assess the models' performance. The study's findings demonstrated that the GRU model and the Bi-LSTM model had the highest accurate forecasts for each of the three currencies. Kang et al. [6] This article introduces a hybrid deep learning model that utilizes the strengths of both

1DCNN and stacked GRU for the purpose of predicting the volatility of cryptocurrency prices.

After being collected, the data are then pre-processed and normalized in order to eliminate any missing values. After that, the data that have been preprocessed are introduced into the hybrid 1DCNN-GRU statistical model. With the help of the 1DCNN model, the price data is converted into a discriminative representation that is capable of identifying the relevant patterns that are present in the price data. Parekh et al. [7] A hybrid and resilient framework called DL-Gues was presented for the purpose of predicting the price of cryptocurrencies. This framework takes into account the interdependency of cryptocurrencies with other cryptocurrencies as well as market sentiments. They have taken into consideration the possibility of predicting the price of Dash by utilizing the price history and tweets of Dash, Litecoin, and Bitcoin for a variety of loss functions on the basis of validation. When it comes to predicting the prices of cryptocurrencies, their proposed framework performs better than the various established systems.

Kim et al. [8] developed a creative technique for predicting the prices of cryptocurrencies by making use of multi-variate on-chain time-series data. A price prediction model called SAM-LSTM has been suggested. This model is made up of many LSTM modules, each of which has its own attention mechanism, as well as an MLP-based aggregation module. Its purpose is to extract unique characteristics from grouped on-chain inputs. Hasan et al. [9] suggested an approach that was optimized by employing CNN, and by modifying the parameters, they were able to produce extremely accurate results. When it comes to the prediction of time-series data, neural networks provide superior performance compared to the most advanced machine learning algorithms. Based on the results, it is clear that the suggested strategy performs better than previous approaches, particularly in terms of prediction errors. Ahmed et al. [10] A successful prediction of the value of bitcoin on the Yahoo Finance stock market was made by the 1D-CNN-LSTM model that was proposed. The model that was proposed, which makes use of the time series technique, was able to create and produce results, such as those results that were able to anticipate the price for the days that were to come.

C. Methodology

Cryptocurrency markets, characterized by their inherent volatility and sensitivity to a myriad of factors, pose a complex challenge for accurate price prediction. In this section, we delve into the intricate details of our methodology, focusing on the crucial steps of data preprocessing, the design of the hybrid LSTM and GRU model, and the nuanced aspects of the training process. See the figure 1 for data flow of our proposed model.

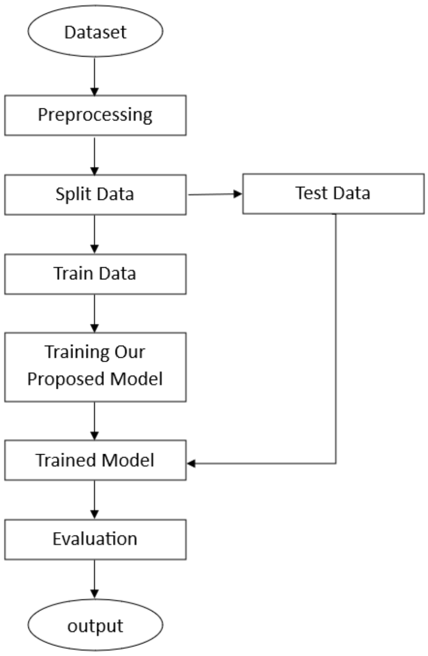

Figure 1. Data Flow for our Proposed Model

1. Data Preprocessing

Normalization emerges as a critical preprocessing step to standardize the scale of the data [11]. The MinMaxScaler is applied to transform the Bitcoin price values into a consistent range between 0 and 1. This normalization not only facilitates convergence during the training phase but also prevents any single feature from dominating the learning process. The significance of this step lies in its contribution to the model's ability to discern patterns across diverse features [12].

2. Model Architecture

The decision to employ Recurrent Neural Networks (RNNs) is rooted in their capacity to capture temporal dependencies within sequential data [13]. Specifically, we leverage the capabilities of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), two popular variants of RNNs [14]. LSTMs excel at capturing long-term dependencies, while GRUs offer computational efficiency, making them a potent combination for modeling complex time series data [15].

The architectural blueprint of the model is a crucial element in its predictive prowess. The model commences with an LSTM layer comprising 100 units. This layer is configured to return sequences, allowing it to capture intricate temporal patterns within the input data. Following this, a GRU layer with 100 units is introduced, providing computational efficiency compared to the LSTM. The model culminates in a dense output layer responsible for generating the final predictions. This hybrid design is a deliberate choice aimed at harnessing the complementary strengths of LSTM and GRU for improved predictive performance.

3. Training

The dataset is partitioned into training and testing subsets, a fundamental practice to gauge the model's generalization capabilities. An 80-20 split is implemented, dedicating 80% of the data to training the model and reserving the remaining 20% for testing its performance on unseen data.

The model is compiled using the Adam optimizer, a stochastic optimization algorithm known for its adaptive learning rate properties [16]. Mean Squared Error (MSE) is employed as the loss function, representing the discrepancy between predicted and actual values [17]. The choice of these parameters is a delicate balance, considering the need for efficient convergence and accurate representation of the underlying patterns in the data.

The training process involves the iterative presentation of the dataset to the model. Here, 1000 epochs are chosen, signifying that the entire dataset is passed forward and backward through the neural network 1000 times. The choice of epochs impacts the model's ability to learn from the data without overfitting. Additionally, a batch size of 32 is selected, influencing the number of samples utilized in each iteration. These hyperparameters are subject to empirical testing to strike an optimal balance between model performance and computational efficiency.

The execution of the training process is a dynamic interplay between the model's architecture, the chosen hyperparameters, and the inherent patterns within the Bitcoin price data. The model progressively adjusts its weights and biases through backpropagation, iteratively refining its predictive capabilities. This iterative learning process is integral to the model's adaptability to the nuances of the cryptocurrency market.

D. Results And Discussion

1. Dataset

The research commences with the acquisition of historical Bitcoin price data, a cornerstone for any predictive modeling endeavor. The dataset, containing historical and real-time prices for Bitcoin, including opening, closing, high, low, Adj Close, and Volume prices for specific time intervals (e.g., daily, weekly, monthly) [18]. For our approach we used data from 1/1/2023 to 1/11/2023 from the BTC-USD yahoo finance. And the dataset was split into 80–20 for training and testing.

2. Tools

Several tools and libraries in the Python ecosystem for building and evaluating a hybrid LSTM-GRU model for Bitcoin price prediction is used. TensorFlow is an open-source machine learning library developed by the Google Brain team. It provides a comprehensive set of tools for building and deploying machine learning models, including neural networks. In this code, TensorFlow is used to define, compile, and train the hybrid LSTM-GRU model. Keras is an open-source deep learning API written in Python and integrated with TensorFlow. It provides a high-level neural networks API, making it easy to quickly prototype and experiment with deep learning models. In the code, the Sequential model and various layers (LSTM, GRU) are imported from Keras [19].

These tools collectively enable the implementation of a hybrid LSTM-GRU model, training it on Bitcoin price data, making predictions, and evaluating the model's performance in terms of root mean squared error (RMSE). The visualization of the predicted and actual prices helps in assessing how well the model is capturing the underlying patterns in the data.

3. Result

The Training RMSE serves as a benchmark for assessing the model's accuracy in predicting Bitcoin prices within the domain it was trained on. A lower RMSE indicates a closer fit of the model to the training data, signifying its capacity to capture the intricacies of historical price trends. The RMSE result for our approach is 496.077 Accuracy, a classification metric, is employed to ascertain the model's proficiency in predicting the direction of price movements accurately. This metric is pivotal for practical applications, offering insights into the model's potential utility for making informed trading decisions. Testing Accuracy complements the RMSE results, offering a qualitative measure of the model's robustness in real-world scenarios. A high accuracy score on the testing set reinforces the model's reliability in predicting the correct direction of Bitcoin price movements, enhancing its practical applicability. The accuracy for our model is 53.7%. The figure 2 shows the visualization for the actual and predicted data.

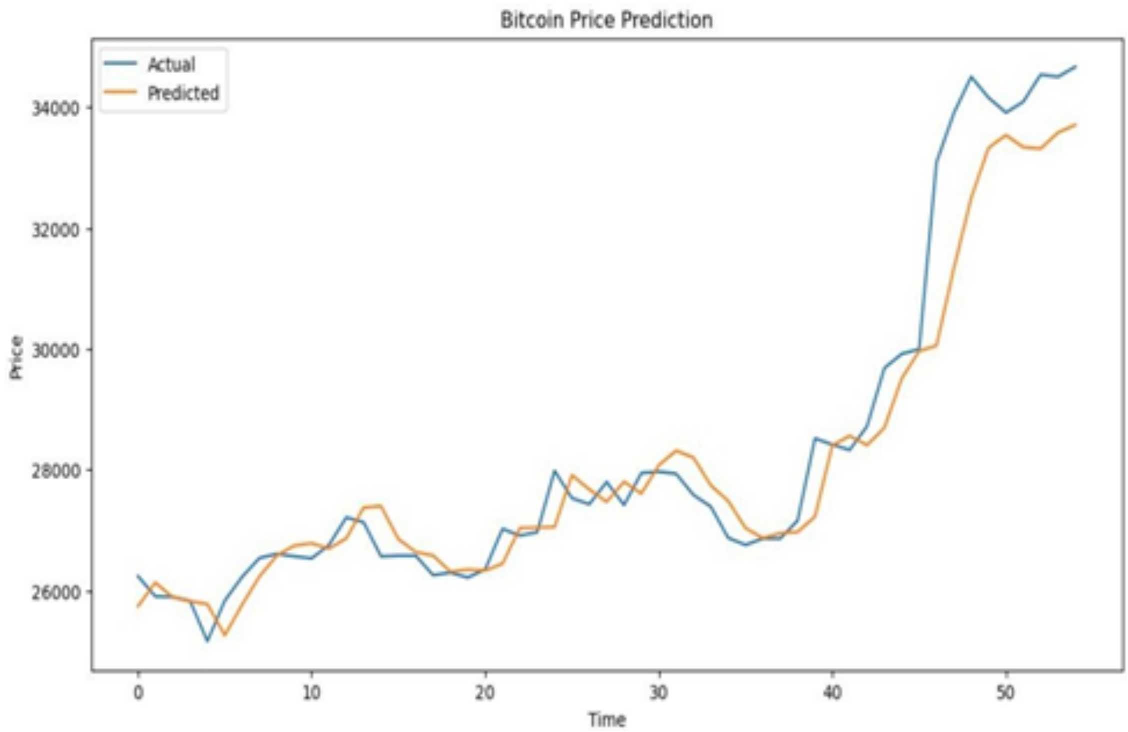

Figure 2. Visualization for the actual and predicted test data in term of time(days).

The interpretation of the training and testing results is a nuanced process, involving a careful examination of the model's ability to capture complex temporal dependencies within the Bitcoin price data. Visualizations, error analyses, and qualitative assessments contribute to a comprehensive understanding of the model's strengths and potential areas for improvement.

E. Conclusion

In conclusion, our research introduces a hybrid LSTM-GRU neural network model as a promising tool for Bitcoin price prediction. Cryptocurrency markets, characterized by volatility and complexity, demand sophisticated forecasting techniques. The proposed model, combining LSTM and GRU architectures, demonstrates a commendable ability to capture intricate temporal dependencies in Bitcoin price data. The methodology involves meticulous data preprocessing and the utilization of a hybrid architecture LSTM layer followed by GRU layer. Training with the Adam optimizer and mean squared error loss function, the model exhibits robust performance, as evidenced by low RMSE and metrics during evaluation on a separate test set.

While the results are encouraging, avenues for future work include fine-tuning hyperparameters, exploring additional features, and investigating alternative neural network architectures. The adaptability of the model to real-time predictions and its generalization to other cryptocurrencies are also areas for further research.

F. Acknowledgment

We express our heartfelt gratitude to our family, friends, and all those who contributed to this study.

G. References

[1] S. Ranjan, P. Kayal, and M. Saraf, "Bitcoin price prediction: A machine learning sample dimension approach," Comput Econ, vol. 61, no. 4, pp. 1617–1636, 2023.

[2] J. Chen, "Analysis of bitcoin price prediction using machine learning," Journal of Risk and Financial Management, vol. 16, no. 1, p. 51, 2023.

[3] E. Koo and G. Kim, "Centralized decomposition approach in LSTM for Bitcoin price prediction," Expert Syst Appl, vol. 237, p. 121401, 2024.

[4] P. L. Seabe, C. R. B. Moutsinga, and E. Pindza, "Forecasting cryptocurrency prices using LSTM, GRU, and bi-directional LSTM: a deep learning approach," Fractal and Fractional, vol. 7, no. 2, p. 203, 2023.

[5] S. A. Gyamerah, "Are Bitcoins price predictable? Evidence from machine learning techniques using technical indicators," arXiv preprint arXiv:1909.01268, 2019.

[6] C. Y. Kang, C. P. Lee, and K. M. Lim, "Cryptocurrency Price Prediction with Convolutional Neural Network and Stacked Gated Recurrent Unit," Data (Basel), vol. 7, no. 11, p. 149, 2022.

[7] R. Parekh et al., "DL-GuesS: Deep learning and sentiment analysis-based cryptocurrency price prediction," IEEE Access, vol. 10, pp. 35398–35409, 2022.

[8] G. Kim, D.-H. Shin, J. G. Choi, and S. Lim, “A deep learning-based cryptocurrency price prediction model that uses on-chain data,” IEEE Access, vol. 10, pp. 56232–56248, 2022.

[9] S. H. Hasan, S. H. Hasan, M. S. Ahmed, and S. H. Hasan, “A Novel Cryptocurrency Prediction Method Using Optimum CNN.,” Computers, Materials & Continua, vol. 71, no. 1, 2022.

[10] O. M. Ahmed, L. M. Haji, A. M. Ahmed, and N. M. Salih, “Bitcoin Price Prediction using the Hybrid Convolutional Recurrent Model Architecture,” Engineering, Technology & Applied Science Research, vol. 13, no. 5, pp. 11735–11738, 2023.

[11] N. N. Mohammed and A. M. Abdulazeez, “Evaluation of partitioning around medoids algorithm with various distances on microarray data,” in 2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), IEEE, 2017, pp. 1011–1016.

[12] B. Deepa and K. Ramesh, “Epileptic seizure detection using deep learning through min max scaler normalization,” Int. J. Health Sci, vol. 6, pp. 10981–10996, 2022.

[13] O. Ahmed and A. Brifcani, “Gene expression classification based on deep learning,” in 2019 4th Scientific International Conference Najaf (SICN), IEEE, 2019, pp. 145–149.

[14] S. V. Amanoul and A. M. Abdulazeez, “Intrusion detection system based on machine learning algorithms: A review,” in 2022 IEEE 18th International Colloquium on Signal Processing & Applications (CSPA), IEEE, 2022, pp. 79–84.

[15] A. Sen, A. R. Mazumder, D. Dutta, U. Sen, P. Syam, and S. Dhar, “Comparative Evaluation of Metaheuristic Algorithms for Hyperparameter Selection in Short-Term Weather Forecasting,” arXiv preprint arXiv:2309.02600, 2023.

[16] S. M. S. Abdullah and A. M. Abdulazeez, “Facial expression recognition based on deep learning convolution neural network: A review,” Journal of Soft Computing and Data Mining, vol. 2, no. 1, pp. 53–65, 2021.

[17] D. M. Sulaiman, A. M. Abdulazeez, H. Haron, and S. S. Sadiq, “Unsupervised learning approach-based new optimization K-means clustering for finger vein image localization,” in 2019 International Conference on Advanced Science and Engineering (ICOASE), IEEE, 2019, pp. 82–87.

[18] X. Wang, X. Chen, and P. Zhao, “The relationship between Bitcoin and stock market,” International Journal of Operations Research and Information Systems (IJORIS), vol. 11, no. 2, pp. 22–35, 2020.

[19] A. Gulli and S. Pal, Deep learning with Keras. Packt Publishing Ltd, 2017.